

修 士 論 文 の 和 文 要 旨

研究科・専攻	大学院 情報システム学 研究科 情報システム設計学専攻 博士前期課程		
氏 名	山崎 敬広	学籍番号	0650037
論 文 題 目	ベイジアンネットワークによる大規模データに対する 協調フィルタリング		
<p>要 旨</p> <p>近年、様々な情報からこのユーザに適したアイテムを自動的に推薦するシステムが注目されている。推薦システムでは、コンテンツ・ベースフィルタリングと、協調フィルタリングの二つの手法が多く用いられてきたが、これらの手法は予測精度が重要視されてきた。Brees ら(1988)は様々な協調フィルタリングの手法を比較し、その結果、ベイジアン・ネットワークが最も高い予測精度を持つことを示した。しかし、Brees らはベイジアン・ネットワークの学習アルゴリズムに BDe と呼ばれる学習アルゴリズムのみを適用している。ベイジアン・ネットワークの予測精度は学習アルゴリズムによって全く異なることが知られており、より優れた予測精度を示す学習アルゴリズムがあるとも考えられる。また、彼らは小規模なデータへのベイジアン・ネットワークの適用しか行っていないため、大規模データに対してベイジアン・ネットワークが有効に適用できるかどうか、わかっていない。</p> <p>本研究では、膨大なページ数を持つ大規模 Web サイトを対象とするベイジアン・ネットワークを用いた推薦システムを開発する。まず、推薦システムに最適なベイジアン・ネットワークの学習アルゴリズムを明らかにするために、四種類の推測統計的手法 (UPSM, BDeu, AIC, MDL) と二種類の記述統計的手法 (MWST, Polytree-MWST) の学習アルゴリズムについて、予測精度比較実験を行った。その結果、一致性を持たない記述統計的手法である MWST, Polytree-MWST が優れた予測精度を持つことがわかった。また、推測統計的手法の学習アルゴリズムの予測精度が MWST, Polytree-MWST に比べて劣ることは、ベイジアン・ネットワークのデータに対する対数尤度の性質によるものであることを示した。</p> <p>次に、大規模データに対応させるために、MWST を元にした高速化アルゴリズムを提案する。MWST は同時エントロピーを用いてネットワークを学習する。そこで、同時エントロピーを求めるために同時確率を求める必要があるが、提案アルゴリズムでは、MWST において二変数の同時確率を求めるのに相関ルールで知られているアプリアリ・アルゴリズムを組み込み高速化を図る。これにより、計算時間が大幅に短縮され、加えて提案アルゴリズムが十分な予測精度を持つことを確認し、その有効性を示した。</p>			